Math 615: Lecture of January 17, 2007

The notion of Gröbner basis is non-trivial and of some interest even when there are no indeterminates, i.e., when $R = K$ is a field, and $F = K^s$.

Consider an $r \times s$ matrix $A = \bigl(a_{i,j}\bigr)$ over a field K. The leftmost nonzero entry of a nonzero row of A is called the *leading* or *initial* entry. Recall that A is said to be in *reduced row echelon form* if it satisfies the following conditions:

(1) The nonzero rows precede the rows that are 0.

(2) The leading entry of every nonzero row is 1.

(3) If there are ρ nonzero rows, and if the leading entry of the i th row in the j_i th column, then $j_1 < j_2 < \cdots < j_\rho$.

(4) If the leading entry of the i th row occurs in the j_i th column, then all other entries in the j_i th column are 0.

The key result from elementary linear algebra about reduced row echelon form is that every $r\times s$ matrix over K has the same row space as a *unique* matrix in reduced row echelon form. Moreover, the given matrix can be put into reduced row echelon form by a sequence of elementary row operations (i.e., multiplying a row by a nonzero scalar, permuting the rows, and adding a multiple of one row to another). This gives a canonical basis for the row space of the original matrix (but this canonical basis does depend on having made a choice of basis for K^s).

Suppose that A is in reduced row echelon form and call the nonzero rows $f_1, \ldots, f_\rho$. The initial term of f_i is e_{j_i}. Condition (4) guarantees that the initial term of f_i does not divide any term in any other f_j. If we have any nonzero element $c_1 f_1 + \cdots + c_\rho f_\rho$ of the row space, its initial term will be $c_i e_{j_i}$ for the smallest value of i such that $c_i \neq 0$. Consequently:

Proposition. *Let the monomial order for K^s be such that $e_1 > e_2 > \cdots > e_s$. An $r \times s$ matrix over the field K is in reduced row echelon form if and only if its nonzero rows precede its zero rows and its nonzero rows form a reduced Gröbner basis for its row space.* □

Relations on monomials and terms

Let $M \subseteq F = R^s$ be any monomial submodule. Since M is generated by monomials $\mu_i e_j$, it follows that

$$M = I_1 e_1 \oplus \cdots \oplus I_s e_s$$

where every I_j is a monomial ideal of R. Understanding the relations on generators for M is therefore equivalent to understanding the relations on generators for several separate monomial ideals of R.

We therefore focus first on understanding generators for the module of relations on a sequence $\mu_1, \ldots, \mu_r$ of monomials in the polynomial ring $R = K[x_1, \ldots, x_n]$. For each pair of monomials μ_i and μ_j with $i \neq j$, we get one "obvious" minimal relation: it comes from the trivial relation that corresponds to the equation

$$\mu_j \mu_i - \mu_i \mu_j = 0$$

by dividing both coefficients μ_j and $-\mu_i$ by $\Delta_{ij} = \mathrm{GCD}(\mu_i, \mu_j)$. (Trivial relatons are also called *Koszul* relations, and the relation obtained by dividing by Δ is sometimes called a *divided Koszul relation.*) Thus, if $I = (\mu_1, \ldots, \mu_r)R$ and we map $R^r \twoheadrightarrow R$ by the map that sends $e_i \mapsto \mu_i$, the kernel will contain the elements

$$\theta_{ij} = \frac{\mu_j}{\Delta_{ij}} e_i - \frac{\mu_j}{\Delta_{ij}} e_j \in R^r.$$

In fact, all relations on μ_i and μ_j are multiples of θ_{ij}. This is a consequence of the following:

Lemma. *Let μ_1 and μ_2 be any two nonzero elements of a UFD R, and let $\Delta = GCD(\mu_1, \mu_2)$, so that $\mu_1 = f_1\Delta$ and and $\mu_2 = f_2\Delta$, with $GCD(f_1, f_2) = 1$. Then $(f_2, -f_1)$ generates the module of relations on μ_1 and μ_2. In other words, if $g_1\mu_1 + g_2\mu_2 = 0$, then (g_1, g_2) is a multiple of $(f_2, -f_1)$.*

Proof. Since $g_1\mu_1 + g_2\mu_2 = 0$, we have that $g_1\Delta f_1 + g_2\Delta f_2 = 0$, and so $g_1 f_1 + g_2 f_2 = 0$. Since f_2 divides $g_1 f_1$ while $\mathrm{GCD}(f_1, f_2) = 1$, we have that f_2 divides g_1, say $g_1 = q f_2$. Then $q f_2 f_1 + g_2 f_2 = 0$, and so $g_2 = -q f_1$, i.e., $(g_1, g_2) = q(f_2, -f_1)$, as required. □

Example. If $\mu_1 = x_1^2 x_2^3 x_3^5 x_4$ and $\mu_2 = x_1^3 x_3^2 x_4^4$, then the trivial or Koszul relation on these two monomials is given by $(x_1^3 x_3^2 x_4^4, -x_1^2 x_2^3 x_3^5 x_4)$, while $\theta_{1,2}$ is the result of factoring out the GCD, which is $x_1^2 x_3^2 x_4$, i.e., $\theta_{12} = (x_1 x_4^3, -x_2^3 x_3^3)$.

We next want to show that the θ_{ij} generate all relations on the μ_j. We first discuss the notion of an $\mathbb{N}^n$-grading, and more general gradings. Let H be a commutative semigroup (which means that the operation is associative) with identity 0, and suppose that the binary operation for H is written additively. An H-graded ring is a ring R with a direct sum decomposition $R = \bigoplus_{h \in H} R_h$ as an abelian group such that $1 \in R_0$ and $R_h R_{h'} \subseteq R_{h+h'}$ for all $h, h' \in H$. An H-graded module M over an H-graded ring R is then an R-module M with a direct sum decomposition $M = \bigoplus_{h \in H} M_h$ as an abelian group such that that $R_h M_{h'} \subseteq M_{h+h'}$ for all $h, h' \in H$. Note that this implies that every M_h is an R_0-module. An element of R or M is called *homogeneous* or a *form* if it is in one of the R_h or M_h.

If f is in an H-graded ring or module, the direct sum decomposition provides a decomposition of f into homogeneous components, one for every element of H, just as in the $\mathbb{N}$-graded case.

An ideal (respectively, a submodule) of an H-graded ring R (respectively, an H-graded module M) is called a *homogeneous* or *graded* ideal (respectively, submodule) if the following two equivalent conditions hold:

(1) It is generated by homogeneous elements.

(2) It contains all of the homogeneous components of all of its elements.

Suppose that we take $H = \mathbb{N}^n$. Then it is easy to see that the polynomial ring $R = K[x_1, \ldots, x_n]$ is $\mathbb{N}^n$-graded, where, if $\alpha \in \mathbb{N}^n$, $R_\alpha = Kx^\alpha$. This is simply a consequence of the fact that $x^\alpha x^\beta = x^{\alpha+\beta}$. In this case, the homogeneous ideals (with respect to the $\mathbb{N}^n$-grading) are precisely the monomial ideals. We can now prove:

Proposition. *The relations θ_{ij} generate all the relations on the monomials $\mu_1, \ldots, \mu_r$.*

Proof. Suppose that we have a relation corresponding to

$$(*) \qquad \sum_{k=1}^{r} f_j \mu_j = 0.$$

(Officially, the relation is $\sum_{j=1}^{r} f_j e_j \in R^r$.) Only finitely many degrees occur when we expanded out all the products occurring in the summation: call these degrees $\alpha_1, \ldots, \alpha_t$. Fix one of these degrees $\alpha_i \in \mathbb{N}^n$. For every i, the the sum of the terms of degree α_i occurring in $(*)$ is 0. If the degree of $\mu_j = \beta_j$, this sum can be represented as

$$(*_i) \qquad \sum_{k=1}^{r} [f_j]_{\alpha_i - \beta_j} \mu_j = 0,$$

where $[f_j]_\gamma$ denotes the degree γ component of f_j. The original relation is the sum of the relations corresponding to the equations $(*_i)$. Therefore, it suffices to show that each of the relations corresponding to one of the equations $(*_i)$ is an R-linear combination of the θ_{ij}. Thus, we need only consider relations in which the degree of every product is x^α for some fixed α. These are *homogeneous* relations.

We may drop the terms with coefficient 0. After renumbering the monomials, we may assume without loss of generality that for every j, f_j is a nonzero term $c_j \mu'_j$ where $\mu'_j \mu_j = x^\alpha$, and α is independent of j. The fact that

$$(**) \qquad \sum_{j=1}^{r} (c_j \mu'_j) \mu_j = 0$$

is then simply the assertion that $\sum_{j=1}^{r} c_j = 0$, and so $c_r = -\sum_{j=1}^{r-1} c_j$.

The given relation is therefore the sum of $r-1$ relations corresponding to equations of the form

$$(***) \qquad c_j \mu'_j \mu_j - c_j \mu'_r \mu_r = 0$$

where $1 \le i \le r-1$. Since this equation corresponds to a relation on just two monomials, namely, μ_j and μ_r, by the preceding Lemma the corresponding relation must be a multiple of θ_{jr}. $\square$

Example. The θ_{ij} are not necessarily a minimal set of generators for the relations on the μ_j. For example, suppose that $\mu_1 = x_2x_3$, $\mu_2 = x_1x_3$ and $\mu_3 = x_1x_2$. Then we have that $\theta_{12} = (x_1,\, x_2,\, 0)$, $\theta_{13} = (x_1,\, 0,\, -x_3)$, and $\theta_{23} = (0,\, x_2,\, -x_3)$. Since $\theta_{13} = \theta_{12} + \theta_{23}$, we only need two of these three relations in a minimal basis.

We want to extend this type of relation θ_{ij} to terms $\gamma_i = c_i\mu_i e_m$ and $\gamma_j = c_j\mu_j e_m$ in a free module F with ordered basis provided that γ_i and γ_j involve the *same* ordered basis element e_m. Here, c_i and c_j are nonzero scalars in K, while $\mathfrak{m}_i$ and μ_j are monomials in R. In this case we let $\Delta_{ij} = \mathrm{GCD}(\gamma_i,\, \gamma_j)$, which we define to be $\mathrm{GCD}(\mu_i, \mu_j)e_m$. We also define γ_i/Δ_{ij} to be $c_i\mu_i/\mathrm{GCD}(\mu_i,\, \mu_j)$, which is a term in R. We still have

$$(\gamma_i/\Delta_{ij})\Delta_{ij} = \gamma_i.$$

We can now define

$$\theta_{ij} = \frac{\gamma_j}{\Delta_{ij}}e_i - \frac{\gamma_j}{\Delta_{ij}}e_j \in R^r,$$

just as in the case of monomials in R. We have at once:

Lemma. *if $\gamma_1, \ldots, \gamma_r$ are terms in F, the module of relations on the elements $\gamma_1, \ldots, \gamma_r$ is generated by the relations θ_{ij} for those choices of i, j such that γ_i and γ_j involve the same element of the ordered basis for F.* □

We shall later discuss a similar result that gives an entire finite free resolution for monomial ideals and submodules. This resolution was discovered by Diana Taylor in the 1960s. However, it is *not minimal.* In fact, the minimal resolution of a monomial ideal may depend on the characteristic of the field K.

The Buchberger criterion and algorithm

Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over a field K, let F be a finitely generated free module with ordered basis, and let $M \subseteq F$ be a submodule. Let $g_1, \ldots, g_r \in M$ be elements that generate M. The following theorem gives necessary and sufficient conditions for the g_j to be a Gröbner basis for M. Once this result is known, one immediately gets an algorithm for enlarging a given set of generators of M to a Gröbner basis for M.

The idea underlying the criterion is to try to produce new elements of $\mathrm{in}(M)$ from the given g_j in an obvious way: first take an efficient monomial linear combination of g_i and g_j that gets their initial terms to cancel. Divide the result with respect to the $g_1, \ldots, g_r$. If the remainder is nonzero, its initial term cannot be in the R-span of the $\mathrm{in}(g_j)$, and we have taken a further step towards finding a Gröbner basis. If all of the remainders are 0, we hope that we already have a Gröbner basis. This is true.

Here is a precise formulation. Let $g_1, \ldots, g_r$ be generators for $M \subseteq F$ and let $\nu_1, \ldots, \nu_r$ be their respective initial terms. For every pair of indices $i \neq j$ such that ν_i and ν_j involve the same element of the ordered basis for F, let

$$G_{ij} = \frac{\nu_j}{\Delta_{ij}}g_i - \frac{\nu_j}{\Delta_{ij}}g_j,$$

where $\Delta_{ij} = \mathrm{GCD}(\nu_i, \nu_j)$. Let h_{ij} be a remainder for division of G_{ij} with respect to $g_1, \dots, g_r$ (the remainder in any standard expression can be used: this need not be the result of using the deterministic division algorithm).

Theorem (Buchberger Criterion). *With notation as in the paragraph above, $g_1, \dots, g_r$ is a Gröbner basis for M if and only if all of the $h_{ij} = 0$.*

We postpone the proof of the sufficiency of the condition momentarily. When we do give the proof, we shall establish a somewhat weaker sufficient condition.

The condition given above is clearly necessary: if the $g_1, \dots, g_r$ form a Gröbner basis for M, then since G_{ij} is clearly in M, our test for membership in M implies that the remainder in any standard expression when we divide G_{ij} with respect to the Gröbner basis $g_1, \dots, g_r$ is 0.

We note that this gives an effective algorithm for finding a Gröbner basis for M given generators $g_1, \dots, g_r$. We calculate values for the h_{ij}. If one of these is nonzero, its initial term cannot be in the span of $\mathrm{in}(g_1), \dots, \mathrm{in}(g_r)$. (To make the process choice-free, we can use the least value of i for which $h_{ij} \neq 0$, and, for that i, the least value of j.) We then enlarge the original set of generators by including this element h_{ij}. The R-span of the initial terms has increased. Since F is Noetherian, the process must terminate, i.e., eventually we reach a set of generators for which all of the h_{ij} are 0. The Buchberger criterion now implies that we have a Gröbner basis for M. This method is called the *Buchberger algorithm.*

We do not, however, have an *a priori* estimate for how many steps will be needed to find the Gröbner basis. In worst cases, the number of steps is double exponential. However, in practice, the method is useful in many of the examples that come up.